



## Modelling of fasting glucose-insulin dynamics from sparse data

Aradóttir, Tinna Björk; Boiroux, Dimitri; Bengtsson, Henrik; Poulsen, Niels Kjølstad

*Published in:*

Proceedings of 40th International Conference of the IEEE Engineering in Medicine and Biology Society

*Link to article, DOI:*

[10.1109/EMBC.2018.8512792](https://doi.org/10.1109/EMBC.2018.8512792)

*Publication date:*

2018

*Document Version*

Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*

Aradóttir, T. B., Boiroux, D., Bengtsson, H., & Poulsen, N. K. (2018). Modelling of fasting glucose-insulin dynamics from sparse data. In *Proceedings of 40th International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 2354-57). IEEE. <https://doi.org/10.1109/EMBC.2018.8512792>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Modelling of fasting glucose-insulin dynamics from sparse data\*

Tinna B. Aradóttir<sup>1,2</sup>, Dimitri Boiroux<sup>2</sup>, Henrik Bengtsson<sup>1</sup>, Niels K. Poulsen<sup>2</sup>

**Abstract**—With the fast growth of diabetes prevalence, the disease is now considered an epidemic. Diabetes is characterized by elevated glucose levels, that may be treated with insulin. Tight control of glucose is essential for prevention of complications and patients' well-being.

In this paper we model the fasting glucose-insulin dynamics in type 2 diabetes, aiming at controlling the glucose level. Relevant clinical data are typically sparse and have a sampling period much greater than the fast dynamics in the glucose-insulin dynamics in humans. We adapt a physiological model such that important slow non-linear dynamics are identifiable and test the resulting model on deterministic simulated data and sparse, slow sampled clinical data.

## I. INTRODUCTION

Prevalence of diabetes has grown extensively over the last years, so much that it is now considered an epidemic. The American Diabetes Association (ADA) predicts that number of diagnosed patients increases from 425 million today to 693 million by 2045 [1]. Although preventable, type 2 diabetes counts for approximately 90% of diabetes cases. The ADA estimates that 20% of health care expenditures in the United States of America are spent on care for people with diabetes [2].

In type 2 diabetes, elevated glucose levels are caused by inadequate production of insulin, reduced response to insulin, or both. High glucose levels can lead to complications in the long term such as eye damage, cardiovascular disease if not treated. However, low glucose levels can cause acute complications, in worst case coma and death. The sequence of treatment intensification in type 2 diabetes starts with lifestyle changes, to oral medication and multiple oral medications, to finally insulin injections [1]. In the current work we focus on the iterative process of initiating long acting insulin treatment of type 2 diabetes.

Response and production of insulin is individual, and therefore finding the right dose of insulin for each patient is important. In standard care, pre-breakfast self-measured blood glucose (SMBG) data are used to adjust dose sizes of long acting insulin. Due to complexity, fear of overdosing, lack of confidence and other factors, the process of finding a sufficient insulin dose can take up to years. A study in the United States of America showed that more than 60%

of type 2 diabetes patients on insulin treatment do not reach recommended treatment goals [3].

In the current work we aim to model fasting glucose in response to long acting insulin. Physiological models of the glucose-insulin regulatory system in healthy and type 1 and type 2 diabetes have been published. These models vary in complexity and purpose, but are most are based on the Bergman minimal model [4], [5]. Some are designed for 24 hour simulations of glucose during meal intake and injections of fast acting insulin in type 1 diabetes [6], [7], [8] or type 2 diabetes [9]. In a paper from 2017, the model by Kanderian et al. [6] was modified to describe fasting glucose in type 2 diabetes in long acting insulin treatment. In this paper the parameters were assessed rather than found by fitting to clinical data [10]. Other groups have investigated models describing progression of type 2 diabetes [11].

In Section II in this paper we outline the problem statement. In Section III, we suggest a physiological model, which parameters can be estimated from clinical data and we briefly describe the methods used in the identification software. In Section IV we describe simulated and clinical data and finally present the results in Section V.

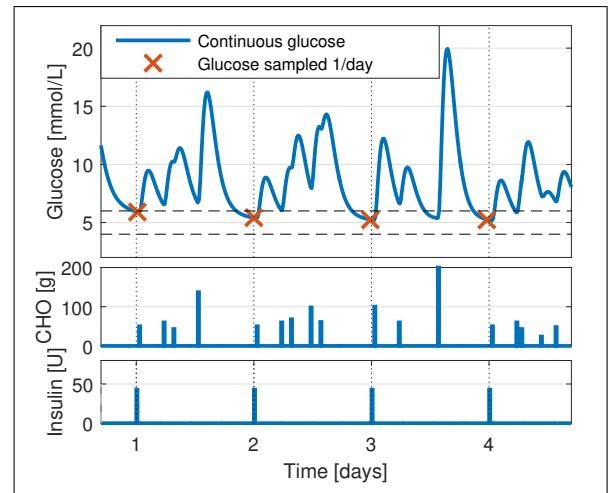


Fig. 1. Simulated glucose data (top) given carbohydrates (middle) and long acting insulin (bottom). The red markers indicate pre-breakfast SMBG. The dashed lines indicate an approximate target range for fasting glucose.

## II. STATEMENT OF THE PROBLEM

The purpose of this work is to create a physiological model of fasting glucose in response to long acting insulin in type 2 diabetes. The aim is to use this model in dose guidance control design in insulin treatment. For result interpretation purposes, such a model should have a physical interpretation.

\*This project is funded by Innovation Fund Denmark through the Industrial PhD project 5189-00033B, and the Danish Diabetes Academy supported by the Novo Nordisk Foundation

<sup>1</sup>Novo Nordisk A/S, DK-2880 Bagsværd, Denmark  
tbao@novonordisk.com, hbss@novonordisk.com

<sup>2</sup>Department of Applied Mathematics and Computer Science, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark  
dibo@dtu.dk, nkpo@dtu.dk

Parameter distributions should also be available for *in silico* simulations of the diverse type 2 diabetes population. Therefore, large scale insulin intensification studies are desirable for parameter estimation.

Parameter estimation in state of the art models of the glucose-insulin regulatory system requires high frequency data, around 1-15 minute sampling time. The top panel of Fig. 1 illustrates in blue a simulation using the model by Kanderian et al. [6], given inputs in the two lower panels. The red markers indicate measuring frequency in most large-scale clinical trials of long acting insulin. Fig. 2 illustrates how data from an insulin intensification study published by Zinman et al. [12] with 763 patients, glucose data was not only registered once per day, but also three out of seven days of the week.

Excitation of physiological systems is limited due to safety. In the case of insulin treatment, large doses of insulin can lead to coma or death, and therefore data below the desired range glucose values are rare.

Given the above limitations, we aim to modify and understand which parameters are identifiable in the state of the art models from the low sampled and sparse clinical data. Ultimately, such a model should predict fasting glucose in insulin intensification treatment of the heterogeneous type 2 diabetes patient group.

### III. METHODS

#### A. A physiological model

We base this work on a model originally published by Kanderian et al. [6] on 24 hour simulations of insulin-glucose dynamics in type 1 diabetes. In [10], this model was augmented with endogenous insulin production and used to simulate fasting glucose in type 2 diabetes. After eliminating the two meal compartments from the model in [10], the

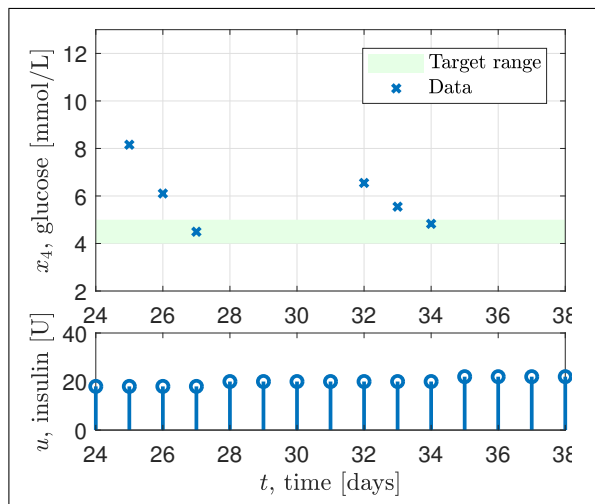


Fig. 2. A cut-out from clinical data for one patient in an insulin intensification trial [12]. The data is sparse with a sampling period of 24 hours.

model is on the form

$$\frac{dx_1}{dt} = \frac{1}{p_1} \frac{u}{p_2} - \frac{1}{p_1} x_1 \quad (1a)$$

$$\frac{dx_2}{dt} = \frac{1}{p_1} x_1 - \frac{1}{p_1} x_2 \quad (1b)$$

$$\frac{dx_3}{dt} = p_3 p_4 (x_2 + p_7 x_4) - p_3 x_3 \quad (1c)$$

$$\frac{dx_4}{dt} = -(p_5 + x_3) x_4 + p_6 \quad (1d)$$

Here,  $u$  is exogenous insulin [U/day] (input variable) and  $x_4$  is glucose concentration in plasma [mmol/L] (controlled variable).  $x_1$  and  $x_2$  denote subcutaneous and plasma insulin concentrations [U/L] where insulin moves with time constant  $p_1$  between the compartments [day].  $x_3$  is insulin effect on glucose [1/day] where  $p_3$  is an inverse time constant describing delay in insulin action following increased insulin concentration in plasma [1/day].  $p_2$  is a gain describing insulin clearance [L/day] and  $p_4$  is a gain describing insulin sensitivity [L/U·day].  $p_5$  is an inverse time constant describing the effect of glucose to eliminate glucose from plasma [1/day] and  $p_6$  is a constant input describing rate of endogenous glucose production [mmol/L·day].  $p_7$  is a parameter added to describe endogenous insulin production, and could be interpreted as glucose sensitivity of the insulin producing cells in the pancreas [U/mmol]. Here insulin production is assumed to increase linearly with fasting glucose.

#### B. Model identifiability

In this paper we work with one observable variable (fasting glucose) and one input variable (insulin). We can therefore not identify both gains in (1),  $p_2$  and  $p_4$ . Setting  $\tilde{x}_1 = x_1 p_2$  [U/day],  $\tilde{x}_2 = x_2 p_2$  [U/day],  $\tilde{x}_3 = x_3 p_2 / p_4$  [U/day],  $\tilde{p}_7 = p_7 p_2$  [U·L/mmol·day] and rewriting (1) gives

$$\frac{d\tilde{x}_1}{dt} = \frac{1}{p_1} u - \frac{1}{p_1} \tilde{x}_1 \quad (2a)$$

$$\frac{d\tilde{x}_2}{dt} = \frac{1}{p_1} \tilde{x}_1 - \frac{1}{p_1} \tilde{x}_2 \quad (2b)$$

$$\frac{d\tilde{x}_3}{dt} = p_3 (\tilde{x}_2 + \tilde{p}_7 x_4) - p_3 \tilde{x}_3 \quad (2c)$$

$$\frac{dx_4}{dt} = -(p_5 + \tilde{p}_4 \tilde{x}_3) x_4 + p_6 \quad (2d)$$

where the gain  $\tilde{p}_4$  is a ratio between the two original gains,  $\tilde{p}_4 = p_4 / p_2$  [1/U].

TABLE I presents mean (sd) parameter values from [6] where units are in L, U, days and mmol (respectively instead of mL and dL, U, min and mg).  $p_1$  is roughly assessed based on knowledge about long acting insulins, and  $p_7$  is calculated such that fasting glucose (for zero input at steady state) in (1) fulfills  $x_{4,ss} = x_{4,0}$ , i.e.

$$p_7 = \frac{1}{p_4 x_{4,0}} \left( \frac{p_6}{x_{4,0}} - p_5 \right) \quad (3)$$

To investigate identifiability of the time constants in (2), we use the values in TABLE I. We compare the values of to the sparse sampling time in the clinical data, i.e. one measurement per day, three days prior to weekly dose

change. The typical value for  $p_3$  in TABLE I is 15.8 day<sup>-1</sup> which corresponds to a time constant of approximately 0.06 days, and we therefore assume we can not estimate it from the clinical data. The assessed value for  $p_1$  is 0.5 days and would be observable from the daily fasting glucose measurements following a change in input. Since the sparse data excludes the first four days following a dose change,  $p_1$  is not observable. The typical value for  $p_5$  is 3.31 day<sup>-1</sup> which is small compared to the term  $\tilde{p}_4 \tilde{x}_3 \approx 15$  day<sup>-1</sup> in the area of interest ( $x_4 \in [4, 6]$ ). We therefore estimate  $\tilde{p}_4$ ,  $\tilde{p}_6$  as well as  $\tilde{p}_7$  from simulated data using (2) using parameters in Table I and clinical data.

### C. Model identification in CTSM-R

CTSM-R, Continuous Time Stochastic Modelling for R, is an open source software for model identification. CTSM-R uses a maximum likelihood approach to identify parameters of an  $l$ -dimensional system of stochastic differential equations on the form

$$d\mathbf{x}_t = f(\mathbf{x}_t, \mathbf{u}_t, t, \theta)dt + \sigma(\mathbf{u}_t, t, \theta)d\mathbf{w}_t \quad (4a)$$

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{u}_k, t_k, \theta) + \mathbf{e}_k \quad (4b)$$

given time series data. Here,  $\theta$  is the set of parameters we want to identify,  $\mathbf{u}_t$  is the input at time  $t$ ,  $\sigma(\mathbf{u}_t, t, \theta)^2$  is the process noise covariance matrix and  $\mathbf{w}_t$  is a Brownian motion path.  $\mathbf{y}_k$  is discrete observations and  $\mathbf{e}_k$  is the measurement error. The likelihood function is a joint probability density, calculated from state and covariance estimates from an extended Kalman filter,

$$L(\theta; \mathcal{Y}_N) = \left( \prod_{k=1}^N p(\mathbf{y}_k | \mathcal{Y}_{k-1}, \theta) \right) p(\mathbf{y}_0, \theta) \quad (5)$$

Here,  $\mathcal{Y}_k = [\mathbf{y}_k, \mathbf{y}_{k-1}, \dots, \mathbf{y}_1, \mathbf{y}_0]$  is a sequence of measurements  $\mathbf{y}_k$ , and the density  $p(\mathbf{y}_k | \mathcal{Y}_{k-1}, \theta)$  is found by

$$p(\mathbf{y}_k | \mathcal{Y}_{k-1}, \theta) = \frac{\exp\left(-\frac{1}{2} \mathbf{e}_k^T \mathbf{R}_{k|k-1}^{-1} \mathbf{e}_k\right)}{\sqrt{\det(\mathbf{R}_{k|k-1})} (\sqrt{2\pi})^l} \quad (6)$$

where  $\hat{\mathbf{y}}_{k|k-1} = E[\mathbf{y}_k | \mathcal{Y}_{k-1}, \theta]$  is an estimate for the measurements,  $\mathbf{R}_{k|k-1} = V[\mathbf{y}_k | \mathcal{Y}_{k-1}, \theta]$  is the covariance of the estimate and  $\mathbf{e}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$  is the innovation.

TABLE I

PARAMETER VALUES FOR THE MODIFIED MODEL (2). MEAN (SD) FOR  $p_2$  THROUGH  $p_6$  ARE MEAN VALUES OF WHAT IS PRESENTED IN [6]. PARAMETER VALUES FOR  $p_1$  IS ASSESSED AND  $p_7$  IS CALCULATED BASED ON OTHER PARAMETERS AND AN INITIAL STATE,  $x_4(0) = 8$  MMOL/L.

|                         | Unit           | Mean (sd)   |
|-------------------------|----------------|-------------|
| $p_1$                   | [day]          | 0.5         |
| $p_2$                   | [L/day]        | 1800 (760)  |
| $p_3$                   | [1/day]        | 15.8 (6.2)  |
| $\tilde{p}_4 = p_4/p_2$ | [1/U]          | 0.44 (0.31) |
| $p_5$                   | [1/day]        | 3.31 (3.17) |
| $p_6$                   | [mmol/L·day]   | 96.7 (63.1) |
| $\tilde{p}_7 = p_7 p_2$ | [U·L/mmol·day] | 2.52        |

CTSM-R allows missing observations and estimates parameters, initial conditions and the measurement and process noise. Other estimates include standard deviance of the estimates and the t-statistic. More information about the methods used in the software is accessible in [13].

## IV. DATA

### A. Simulated data

Simulated data are generated from the model in (2) using the mean values of the parameters in TABLE I. We excite the system with a range of insulin injections such that fasting glucose levels span clinically relevant glucose concentrations; hyperglycemia ( $> 6$  mmol/L), normoglycemia ( $< 6$  mmol/L and  $> 3.9$  mmol/L) and hypoglycemia ( $< 3.9$  mmol/L). The blue markers in Fig. 3 indicate simulated data points.

### B. Clinical data

In this study we use data from a Phase III clinical trial, published in [12]. In this study, 773 adult type 2 diabetes patients initiated insulin degludec treatment on a once daily treatment regimen. Doses were increased weekly for up to a year, based on fasting glucose measurements. Fasting glucose was logged the last three days of the week as illustrated in Fig. 2. In clinical studies, excitation of the system is bounded by safety. The International Hypoglycaemia Study Group defines glucose levels lower than 3.0 mmol/L as clinically significant hypoglycemia and intensification algorithms used in clinical care usually control fasting glucose to a target zone, typically 4.0-5.0 mmol/L or 4.0-6.0 mmol/L [3]. Therefore we can not expect to see glucose values lower than around 3.0-4.0 mmol/L.

## V. RESULTS

### A. Fitting to simulated data

We estimate the set of parameters from simulated data where parameter values are known. We estimate  $\tilde{p}_4$ ,  $p_6$  and  $\tilde{p}_7$  based on the simulated data in Fig. 3. The red line in the figure illustrates the one step prediction, and TABLE II lists the parameter estimates,  $\hat{\theta}$ , true values,  $\theta$  and the initial parameter value  $\theta_0$ . All parameter estimates are significant and close to the true values. We therefore conclude that the three parameters are identifiable, and the objective function (5) seems to be convex.

TABLE II

PARAMETER ESTIMATES AND TRUE VALUES FOR THE SIMULATED DATA.

| $\theta$      | $\theta_0$ | $\theta^*$ | $\hat{\theta}$ | 95% confidence interval* | $p(>  t )$ |
|---------------|------------|------------|----------------|--------------------------|------------|
| $\tilde{p}_4$ | 10         | 0.44       | 0.438          | [0.436, 0.440]           | $< 0.05$   |
| $p_6$         | 500        | 96.7       | 97.1           | [96.6, 97.5]             | $< 0.05$   |
| $\tilde{p}_7$ | 10         | 2.52       | 2.51           | [2.508, 2.511]           | $< 0.05$   |

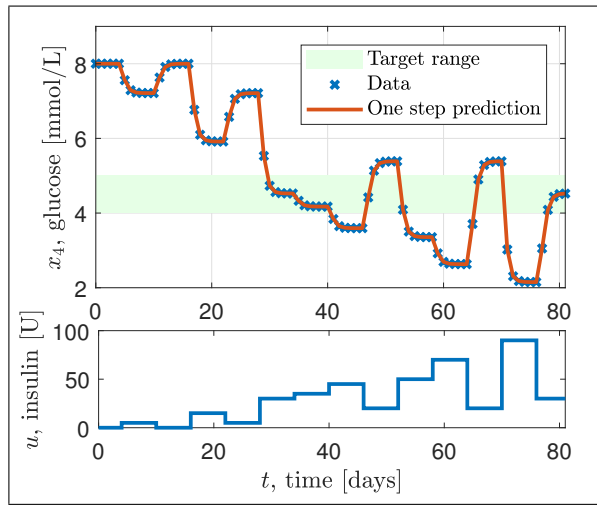


Fig. 3. Simulated glucose data using (1) (top, blue markers) given insulin input (lower panel). The red line shows the one step prediction by the model fit in TABLE II found in CTSM-R.

### B. Fitting to clinical data

Fig. 4 illustrates in blue the glucose and insulin data for one patient in the study published by Zinman et al. [12]. We use CTSM-R to estimate  $\hat{p}_4$ ,  $p_6$  and  $\hat{p}_7$  from these data. The estimates are listed in TABLE III, and the model fit is shown in red in Fig. 4. The estimated autocorrelation function shows that there is no significant correlation between residuals up to a lag of 20, and a sign test for whiteness shows that we can not reject the null-hypothesis that residuals are continuously distributed around zero.

TABLE III  
PARAMETER ESTIMATES FOR THE CLINICAL DATA.

| $\theta$    | $\hat{\theta}$ | 95% confidence interval* | $p(>  t )$ |
|-------------|----------------|--------------------------|------------|
| $\hat{p}_4$ | 1.80           | [1.67, 1.94]             | < 0.05     |
| $p_6$       | 368            | [361, 376]               | < 0.05     |
| $\hat{p}_7$ | 1.68           | [1.36, 2.00]             | < 0.05     |

## VI. CONCLUSION

This work suggests that the slow dynamics of the fasting glucose-insulin system can be described with a modified version of a model of the fast glucose-insulin dynamics. We keep a number of the parameters in the four-compartment model fixed, and thereby identify important parameters describing non-linear dynamics from sparse slow sampled clinical data.

Time constants of the system are however not identifiable due to the sparsity of the data. Design of experiment is an interesting aspect in this context, with respect to data quality, system excitation and sampling frequency.

## REFERENCES

- [1] American Diabetes Association, *Standards of medical Care in Diabetes*, 2017.
- [2] M. Petersen, "Economic costs of diabetes in the U.S. in 2012," *Diabetes Care*, vol. 39, no. 7, pp. 1033–1046, 2016.

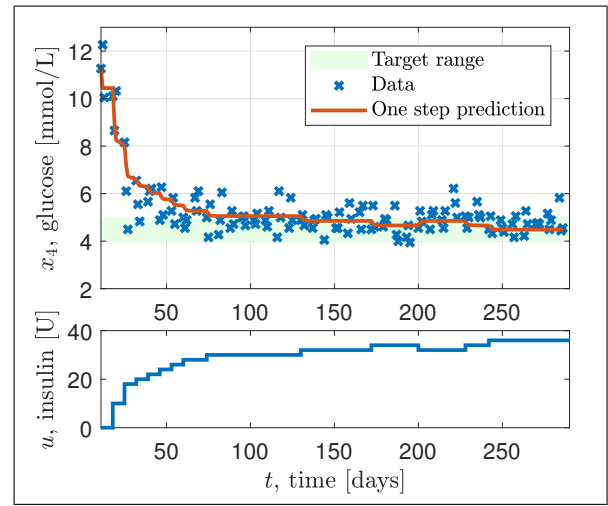


Fig. 4. Clinical glucose data from the long-acting insulin intensification trial [12] (top, blue markers) for one patient, and corresponding insulin data (bottom). The red line indicates the one step prediction by the model fit in TABLE III found in CTSM-R. The green zone in the top panel shows the target range for glucose in the clinical trial.

- [3] S. Arnolds, T. Heise, F. Flacke, and J. Sieber, "Common Standards of Basal Insulin Titration in T2DM," *Journal of Diabetes Science and Technology*, vol. 7, no. 3, pp. 771–788, 2013. [Online]. Available: <http://dst.sagepub.com/content/7/3/771.abstract>
- [4] R. N. Bergman, Y. Z. Ider, C. R. Bowden, and C. Cobelli, "Quantitative estimation of insulin sensitivity," *Am J Physiol*, vol. 23, pp. E667–E677, 1979.
- [5] G. Toffolo, R. N. Bergman, D. T. Finegood, C. R. Bowden, and C. Cobelli, "Quantitative estimation of beta cell sensitivity to glucose in the intact organism: a minimal model of insulin kinetics in the dog," *Diabetes*, vol. 29, pp. 979–990, 1980.
- [6] S. Kanderian, S. Weinzimer, G. Voskanyan, and G. M. Steil, "Identification of intraday metabolic profiles during closed-loop glucose control in individuals with type 1 diabetes," *Journal of diabetes science and technology*, vol. 3, no. 5, pp. 1047–57, 2009.
- [7] R. Hovorka, V. Canonico, L. J. Chassin, U. Haueter, M. Massi-Benedetti, M. O. Federici, T. R. Pieber, H. C. Schaller, L. Schaupp, T. Vering, and M. E. Wilinska, "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes," *Physiological Measurement*, vol. 25, no. 4, pp. 905–920, 2004.
- [8] C. Dalla Man, R. A. Rizza, and C. Cobelli, "Meal simulation model of the glucose- insulin system," *IEEE Trans Biom Engin*, vol. 54, no. 10, pp. 1740–1749, 2007.
- [9] P. M. Jauslin, N. Frey, and M. O. Karlsson, "Modeling of 24-hour glucose and insulin profiles of patients with type 2 diabetes," *Journal of clinical pharmacology*, vol. 51, no. 2, pp. 153–164, 2011.
- [10] T. B. Aradóttir, D. Boiroux, H. Bengtsson, J. Kildegaard, B. V. Orden, and J. B. Jørgensen, "Model for Simulating Fasting Glucose in Type 2 Diabetes and the Effect of Adherence to Treatment," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 15 086–15 091, 2017.
- [11] A. De Gaetano, T. Hardy, B. Beck, E. Abu-Raddad, P. Palumbo, J. Bue-Valleskey, and N. Pørksen, "Mathematical models of diabetes progression," *American Journal of Physiology - Endocrinology and Metabolism*, vol. 295, no. 6, pp. E1462–E1479, 2008.
- [12] B. Zinman, A. Philis-Tsimikas, B. Cariou, Y. Handelsman, H. W. Rodbard, T. Johansen, L. Endahl, and C. Mathieu, "Insulin Degludec Versus Insulin Glargine in Insulin-Naive Patients With Type 2 Diabetes," *Diabetes Care*, vol. 35, no. 12, pp. 2464 LP – 2471, 2012.
- [13] N. R. Kristensen, H. Madsen, and S. B. Jrgensen, "Parameter Estimation in Stochastic Grey-Box Models," *Automatica*, vol. 40, pp. 225–237, 2004.